Diversity Analysis of Symbol-by-Symbol Linear Equalizers

Ali Tajer*   Aria Nosratinia [†]   Naofal Al-Dhahir[†]

Abstract

In frequency-selective channels linear receivers enjoy significantly-reduced complexity compared with maximum likelihood receivers at the cost of performance degradation which can be in the form of a loss of the inherent frequency diversity order or reduced coding gain. This paper demonstrates that the minimum mean-square error symbol-by-symbol linear equalizer incurs no diversity loss compared to the maximum likelihood receivers. In particular, for a channel with memory ν, it achieves the full diversity order of $(\nu + 1)$ while the zero-forcing symbol-by-symbol linear equalizer always achieves a diversity order of one.

I. INTRODUCTION

In broadband wireless communication systems, the coherence bandwidth of the fading channel is significantly less than the transmission bandwidth. This results in inter-symbol interference (ISI) and at the same time provides frequency diversity that can be exploited at the receiver to enhance transmission reliability [1]. It is well-known that for Rayleigh *flat*-fading channels, the error rate decays only linearly with signal-to-noise ratio (SNR) [1]. For frequency-selective channels, however, proper exploitation of the available frequency diversity forces the error probability to decay at a possibly higher rate and, therefore, can potentially achieve higher diversity gains, depending on the detection scheme employed at the receiver.

While maximum likelihood sequence detection (MLSD) [2] achieves optimum performance over ISI channels, its complexity (as measured by the number of MLSD trellis states) grows *exponentially* with the spectral efficiency and the channel memory. As a low-complexity alternative, filtering-based symbol-by-symbol equalizers (both linear and decision feedback) have been widely used over the past four decades (see [3] and [4] for excellent tutorials). Despite their long history and successful commercial deployment, the performance of symbol-by-symbol linear equalizers over wireless fading channels is not fully characterized. More specifically, it is not known whether their observed sub-optimum performance is due to their inability to fully exploit the channel's frequency diversity or due to a degraded performance in

* Electrical Engineering Department, Princeton University, Princeton, NJ 08544.

† Electrical Engineering Department, University of Texas at Dallas, Richardson, TX 75083.

combating the residual inter-symbol interference. Therefore, it is of paramount importance to investigate the frequency diversity order achieved by linear equalizers, which is the subject of this paper. Our analysis shows that while single-carrier infinite-length symbol-by-symbol minimum mean-square error (MMSE) linear equalization achieves full frequency diversity, zero-forcing (ZF) linear equalizers cannot exploit the frequency diversity provided by frequency-selective channels.

A preliminary version of the results of this paper on the MMSE linear equalization has partially appeared in [5] and the proofs available in [5] are skipped and referred to wherever necessary. The current paper provides two key contributions beyond [5]. First, the diversity analysis of ZF equalizers is added. Second, the MMSE analysis in [5] lacked a critical step that was not rigorously complete; the missing parts that have key role in analyzing the diversity order are provided in this paper.

II. SYSTEM DESCRIPTIONS

A. Transmission Model

Consider a quasi-static ISI wireless fading channel with memory length ν and channel impulse response (CIR) denoted by $\boldsymbol{h} = [h_0, \ldots, h_\nu]$. Without loss of generality, we restrict our analyses to CIR realizations with $h_0 \neq 0$. The output of the channel at time k is given by

$$y_k = \sum_{i=0}^{\nu} h_i x_{k-i} + n_k \ , \tag{1}$$

where x_k is the input to the channel at time k satisfying the power constraint $\mathbb{E}[|x_k|^2] \leq P_0$ and n_k is the additive white Gaussian noise term distributed as $\mathcal{N}_{\mathbb{C}}(0, N_0)$[1]. The CIR coefficients $\{h_i\}_{i=0}^{\nu}$ are distributed independently with h_i being distributed as $\mathcal{N}_{\mathbb{C}}(0, \lambda_i)$. Defining the D-transform of the input sequence $\{x_k\}$ as $X(D) = \sum_k x_k D^k$, and similarly defining $Y(D), H(D)$, and $Z(D)$, the baseband input-output model can be cast in the D-domain as $Y(D) = H(D) \cdot X(D) + Z(D)$. The superscript $*$ denotes complex conjugate and we use the shorthand D^{-*} for $(D^{-1})^*$. We define $\mathsf{SNR} \triangleq \frac{P_0}{N_0}$ and say that the functions $f(\mathsf{SNR})$ and $g(\mathsf{SNR})$ are *exponentially equal*, indicated by $f(\mathsf{SNR}) \doteq g(\mathsf{SNR})$, when

$$\lim_{\mathsf{SNR} \to \infty} \frac{\log f(\mathsf{SNR})}{\log \mathsf{SNR}} = \lim_{\mathsf{SNR} \to \infty} \frac{\log g(\mathsf{SNR})}{\log \mathsf{SNR}} \ . \tag{2}$$

The operators $\dot{\leq}$ and $\dot{\geq}$ are defined in a similar fashion. Furthermore, we say that the *exponential order* of $f(\mathsf{SNR})$ is d if $f(\mathsf{SNR}) \doteq \mathsf{SNR}^d$.

B. Linear Equalization

The zero-forcing (ZF) linear equalizers are designed to produce an ISI-free sequence of symbols and ignore the resulting noise enhancement. By taking into account the *combined* effects of the ISI channel and

---

[1]$\mathcal{N}_{\mathbb{C}}(a, b)$ denotes a complex Gaussian distribution with mean a and variance b.

its corresponding matched-filter, the ZF linear equalizer in the D-domain is given by [6, Equation (3.87)]

$$W_{\mathrm{zf}}(D) = \frac{\|\boldsymbol{h}\|}{H(D)H^*(D^{-*})} \; , \tag{3}$$

where the $\|\boldsymbol{h}\|$ is the ℓ_2-norm of $\boldsymbol{h}$, i.e., $\|\boldsymbol{h}\|^2 = \sum_{i=0}^{\nu} |h_i|^2$. The variance of the noise seen at the output of the ZF equalizer is the key factor in the performance of the equalizer and is given by

$$\sigma_{\mathrm{zf}}^2 \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{N_0}{|H(e^{-ju})|^2} \, du \; . \tag{4}$$

Therefore, the decision-point signal-to-noise ratio for any CIR realization $\boldsymbol{h}$ and $\mathsf{SNR} = \frac{P_0}{N_0}$ is

$$\gamma_{\mathrm{zf}}(\mathsf{SNR}, \boldsymbol{h}) \triangleq \mathsf{SNR} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|H(e^{-ju})|^2} \, du \right]^{-1} . \tag{5}$$

MMSE linear equalizers are designed to strike a balance between ISI reduction and noise enhancement through minimizing the combined residual ISI and noise level. Given the combined effect of the ISI channel and its corresponding matched-filter, the MMSE linear equalizer in the D-domain is [6, Equation (3.148)]

$$W_{\mathrm{mmse}}(D) = \frac{\|\boldsymbol{h}\|}{H(D)H^*(D^{-*}) + \mathsf{SNR}^{-1}} \; . \tag{6}$$

The variance of the residual ISI and the noise variance as seen at the output of the equalizer is

$$\sigma_{\mathrm{mmse}}^2 \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{N_0}{|H(e^{-ju})|^2 + \mathsf{SNR}^{-1}} \, du \; . \tag{7}$$

Hence, the *unbiased*[2] decision-point signal-to-noise ratio at for any CIR realization $\boldsymbol{h}$ and SNR is

$$\gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h}) \triangleq \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\mathsf{SNR}|H(e^{-ju})|^2 + 1} \, du \right]^{-1} - 1 \; . \tag{8}$$

C. Diversity Gain

For a transmitter sending information bits at spectral efficiency R bits/sec/Hz, the system is said to be in *outage* if the ISI channel is faded such that it cannot sustain an arbitrarily reliable communication at the intended communication spectral efficiency R, or equivalently, the mutual information $I(x_k, \tilde{y}_k)$ falls below the target spectral efficiency R, where $\tilde{y}_k$ denotes the equalizer output. The probability of such outage for the signal-to-noise ratio $\gamma(\mathsf{SNR}, \boldsymbol{h})$ is

$$P_{\mathrm{out}}(R, \mathsf{SNR}) \triangleq P_{\boldsymbol{h}} \left(\log \left[1 + \gamma(\mathsf{SNR}, \boldsymbol{h}) \right] < R \right) , \tag{9}$$

where the probability is taken over the ensemble of all CIR realizations $\boldsymbol{h}$. The outage probability at high transmission powers ($\mathsf{SNR} \to \infty$) is closely related to the *average pairwise error probability*, denoted by

---

[2]All MMSE equalizers are biased. Removing the bias decreases the decision-point signal-to-noise ratio by 1 (in linear scale) but improves the error probability [7]. All the results provided in this paper are valid for biased receivers as well.

$P_{\text{err}}(R, \text{SNR})$, which is the probability that a transmitted codeword $\boldsymbol{c}_i$ is erroneously detected in favor of another codeword $\boldsymbol{c}_j$, $j \neq i$, i.e.,

$$P_{\text{err}}(R, \text{SNR}) \triangleq \mathbb{E}_{\boldsymbol{h}}\left[P\Big(\boldsymbol{c}_i \to \boldsymbol{c}_j \mid \boldsymbol{h}\Big)\right] . \tag{10}$$

When deploying channel coding with arbitrarily long code-length, the outage and error probabilities decay at the same rate with increasing SNR and have the same exponential order [8] and therefore

$$P_{\text{out}}(R, \text{SNR}) \doteq P_{\text{err}}(R, \text{SNR}) . \tag{11}$$

This is intuitively justified by noting that in high SNR regimes, the effect of channel noise is diminishing and the dominant source of erroneous detection is channel fading which, as mentioned above, is also the source of outage events. As a result, in our setup, diversity order which is the negative of the exponential order of the average pairwise error probability $P_{\text{err}}(R, \text{SNR})$ is computed as

$$d = - \lim_{\text{SNR} \to \infty} \frac{\log P_{\text{out}}(R, \text{SNR})}{\log \text{SNR}} . \tag{12}$$

III. DIVERSITY ORDER OF MMSE LINEAR EQUALIZATION

The main result of this paper for the MMSE linear equalizers is given in the following theorem.

Theorem 1: For an ISI channel with channel memory length $\nu \geq 1$, and symbol-by-symbol MMSE linear equalization we have

$$P_{\text{err}}^{\text{mmse}}(R, \text{SNR}) \doteq \text{SNR}^{-(\nu+1)}.$$

The sketch of the proof is as follows. First, we find a lower bound on the unbiased decision-point signal-to-noise ratio (SNR) and use this lower bound to show that for small enough spectral efficiencies a full diversity order of $(\nu + 1)$ is achievable. The proof of the diversity gain for low spectral efficiencies is offered in Section III-A. In the second step, we show that increasing the spectral efficiency to any arbitrary level does not incur a diversity loss, concluding that MMSE linear equalization is capable of collecting the full frequency diversity order of ISI channels. Such generalization of the results presented in Section III-A to arbitrary spectral efficiencies is analyzed in Section III-B.

A. Full Diversity for Low Spectral Efficiencies

We start by showing that for arbitrarily small data transmission spectral efficiencies, R, full diversity is achievable. Corresponding to each CIR realization $\boldsymbol{h}$, we define the function $f(\boldsymbol{h}, u) \triangleq |H(e^{-ju})|^2 - \|\boldsymbol{h}\|^2$ for which after some simple manipulations we have

$$f(\boldsymbol{h}, u) = \sum_{k=-\nu}^{\nu} c_k\, e^{jku} , \quad \text{where } c_0 = 0, \ c_{-k} = c_k^*, \ c_k = \sum_{m=0}^{\nu-k} h_m h_{m+k}^* \quad \text{for} \quad k \in \{1, \ldots, \nu\} . \tag{13}$$

Therefore, $f(\boldsymbol{h}, u)$ is a trigonometric polynomial of degree ν that is periodic with period 2π and in the open interval $[-\pi, \pi]$ has at most 2ν roots [9]. Corresponding to the CIR realization $\boldsymbol{h}$ we define the set

$$\mathcal{D}(\boldsymbol{h}) \triangleq \{u \in [-\pi, \pi] \ : \ f(\boldsymbol{h}, u) > 0\} \ ,$$

and use the convention $|\mathcal{D}(\boldsymbol{h})|$ to denote the measure of $\mathcal{D}(\boldsymbol{h})$, i.e., the aggregate lengths of the intervals over which $f(\boldsymbol{h}, u)$ is strictly positive. In the following lemma, we obtain a lower bound on $|\mathcal{D}(\boldsymbol{h})|$ which is instrumental in finding a lower bound on $\gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h})$.

Lemma 1: There exists a real number $C > 0$ such that for all non-zero CIR realizations $\boldsymbol{h}$, i.e. $\forall \boldsymbol{h} \neq \boldsymbol{0}$, we have that $|\mathcal{D}(\boldsymbol{h})| \geq C \left(2(2\nu + 1)^3\right)^{-\frac{1}{2}}$.

Proof: According to (13) we immediately have $\int_{-\pi}^{\pi} f(\boldsymbol{h}, u) \ du = 0$. By invoking the definition of $\mathcal{D}(\boldsymbol{h})$ and noting that $[-\pi, \pi] \backslash \mathcal{D}(\boldsymbol{h})$ includes the values of u for which $f(\boldsymbol{h}, u)$ is negative, we have

$$\int_{\mathcal{D}(\boldsymbol{h})} f(\boldsymbol{h}, u) \ du = - \int_{[-\pi, \pi] \backslash \mathcal{D}(\boldsymbol{h})} f(\boldsymbol{h}, u) \ du \quad \Rightarrow \quad \int_{-\pi}^{\pi} |f(\boldsymbol{h}, u)| \ du = 2 \int_{\mathcal{D}(\boldsymbol{h})} f(\boldsymbol{h}, u) \ du \ . \quad (14)$$

Also by noting that $f(\boldsymbol{h}, u) = |H(e^{-ju})|^2 - \|\boldsymbol{h}\|^2$, $f(\boldsymbol{h}, u)$ has clearly a real value for any u. Moreover, by invoking (13) from the Cauchy-Schwartz inequality we obtain

$$f(\boldsymbol{h}, u) \leq |f(\boldsymbol{h}, u)| \leq \left(\sum_{k=-\nu}^{\nu} |c_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k=-\nu}^{\nu} |e^{jku}|^2 \right)^{\frac{1}{2}} = \left(2(2\nu + 1) \sum_{k=1}^{\nu} |c_k|^2 \right)^{\frac{1}{2}} . \quad (15)$$

Equations (14) and (15) together establish that

$$|\mathcal{D}(\boldsymbol{h})| \geq \frac{1}{2} \left(2(2\nu + 1) \sum_{k=1}^{\nu} |c_k|^2 \right)^{-\frac{1}{2}} \int_{-\pi}^{\pi} |f(\boldsymbol{h}, u)| \ du \ . \quad (16)$$

Next we strive to find a lower bound on $\int_{-\pi}^{\pi} |f(\boldsymbol{h}, u)| \ du$, which according to (13) is equivalent to finding a lower bound on the ℓ_1 norm of a sum of exponential terms. Obtaining lower bounds on the ℓ_1 norm of exponential sums has a rich literature in the mathematical analysis and we use a relevant result in this literature that is related to Hardy's inequality [10, Theorem 2].

Theorem 2: *[10, Theorem 2]* There is a real number $C > 0$ such that for any given sequence of increasing integers $\{n_k\}$, and complex numbers $\{d_k\}$, and for any $N \in \mathbb{N}$ we have

$$\int_{-\pi}^{\pi} \left| \sum_{k=1}^{N} d_k \, e^{jn_k u} \right| \ du \geq C \sum_{k=1}^{N} \frac{|d_k|}{k} \ . \quad (17)$$

By setting $N = 2\nu + 1$ and $d_k = c_{k-(\nu+1)}$ and $n_k = k - (\nu + 1)$ for $k \in \{1, \dots, 2\nu + 1\}$ from (17) it is concluded that there exists $C > 0$ that for each set of $\{c_{-\nu}, \dots, c_\nu\}$ we have

$$\int_{-\pi}^{\pi} |f(\boldsymbol{h}, u)| \ du \geq C \sum_{k=1}^{2\nu+1} \frac{|c_{k-(\nu+1)}|}{k} \geq \frac{C}{2\nu + 1} \sum_{k=1}^{2\nu+1} |c_{k-(\nu+1)}| = \frac{2C}{2\nu + 1} \sum_{k=1}^{\nu} |c_k| \ , \quad (18)$$

where the last equality holds by noting that $c_{-k} = c_k^*$ and $c_0 = 0$. Combining (16) and (18) provides

$$|\mathcal{D}(\boldsymbol{h})| \geq C \left(2(2\nu + 1)^3\right)^{-\frac{1}{2}} \underbrace{\frac{\sum_{k=1}^{\nu} |c_k|}{\sqrt{\sum_{k=1}^{\nu} |c_k|^2}}}_{\geq 1} \geq C \left(2(2\nu + 1)^3\right)^{-\frac{1}{2}} \ , \quad (19)$$

which concludes the proof. ∎

Now by using Lemma 1 for any CIR realization $\boldsymbol{h}$ and SNR we find a lower bound on $\gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h})$ that depends on $\boldsymbol{h}$ through $\|\boldsymbol{h}\|$ only. By defining $\mathcal{D}^c(\boldsymbol{h}) = [-\pi, \pi] \backslash \mathcal{D}(\boldsymbol{h})$ we have

$$
1 + \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h}) \overset{(8)}{=} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\mathsf{SNR}|H(e^{ju})|^2 + 1} \, du \right]^{-1} = \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\mathsf{SNR}(f(\boldsymbol{h}, u) + \|\boldsymbol{h}\|^2) + 1} \, du \right]^{-1}
$$

$$
= \left[\frac{1}{2\pi} \int_{\mathcal{D}(\boldsymbol{h})} \frac{1}{\mathsf{SNR}(\underbrace{f(\boldsymbol{h}, u)}_{>0} + \|\boldsymbol{h}\|^2) + 1} \, du + \frac{1}{2\pi} \int_{\mathcal{D}^c(\boldsymbol{h})} \frac{1}{\mathsf{SNR}\underbrace{|H(e^{ju})|^2}_{\geq 0} + 1} \, du \right]^{-1}
$$

$$
\geq \left[\frac{1}{2\pi} \int_{\mathcal{D}(\boldsymbol{h})} \frac{1}{\mathsf{SNR}\|\boldsymbol{h}\|^2 + 1} \, du + \frac{1}{2\pi} \int_{\mathcal{D}^c(\boldsymbol{h})} 1 \, du \right]^{-1}
$$

$$
= \left[\frac{|\mathcal{D}(\boldsymbol{h})|}{2\pi} \cdot \frac{1}{\mathsf{SNR}\|\boldsymbol{h}\|^2 + 1} + \left(1 - \frac{|\mathcal{D}(\boldsymbol{h})|}{2\pi} \right) \right]^{-1}
$$

$$
= \left[1 - \frac{|\mathcal{D}(\boldsymbol{h})|}{2\pi} \left(1 - \frac{1}{\mathsf{SNR}\|\boldsymbol{h}\|^2 + 1} \right) \right]^{-1}
$$

$$
\overset{(19)}{\geq} \left[1 - \frac{C \left(2(2\nu + 1)^3 \right)^{-\frac{1}{2}}}{2\pi} \left(1 - \frac{1}{\mathsf{SNR}\|\boldsymbol{h}\|^2 + 1} \right) \right]^{-1} . \tag{20}
$$

By defining $C' \triangleq \frac{C(2(2\nu+1)^3)^{-\frac{1}{2}}}{2\pi}$, for the outage probability corresponding to the target spectral efficiency R we have

$$
P_{\mathrm{out}}^{\mathrm{mmse}}(R, \mathsf{SNR}) \overset{(9)}{=} P_{\boldsymbol{h}} \left(1 + \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h}) < 2^R \right) \overset{(20)}{\leq} P_{\boldsymbol{h}} \left\{ 1 - C' \left(1 - \frac{1}{\mathsf{SNR}\|\boldsymbol{h}\|^2 + 1} \right) > 2^{-R} \right\}
$$

$$
= P_{\boldsymbol{h}} \left\{ 1 - \frac{1 - 2^{-R}}{C'} < \frac{1}{\mathsf{SNR}\|\boldsymbol{h}\|^2 + 1} \right\} . \tag{21}
$$

If

$$
1 - \frac{1 - 2^{-R}}{C'} > 0 \qquad \text{or equivalently} \qquad R < R_{\max} \triangleq \log_2 \left(\frac{1}{1 - C'} \right) , \tag{22}
$$

then the probability term in (21) can be restated as

$$
P_{\boldsymbol{h}} \left\{ \mathsf{SNR}\|\boldsymbol{h}\|^2 < \frac{1 - 2^{-R}}{C' - (1 - 2^{-R})} \right\} = P_{\boldsymbol{h}} \left\{ \mathsf{SNR}\|\boldsymbol{h}\|^2 < \frac{2^R - 1}{1 - 2^{R - R_{\max}}} \right\} . \tag{23}
$$

Therefore, based on (21)-(23) for all $0 < R < R_{\max}$ we have

$$
P_{\mathrm{out}}^{\mathrm{mmse}}(R, \mathsf{SNR}) \leq P_{\boldsymbol{h}} \left\{ \mathsf{SNR}\|\boldsymbol{h}\|^2 < \frac{2^R - 1}{1 - 2^{R - R_{\max}}} \right\} = P_{\boldsymbol{h}} \left\{ \mathsf{SNR} \sum_{m=0}^{\nu} |h_m|^2 < \frac{2^R - 1}{1 - 2^{R - R_{\max}}} \right\}
$$

$$
\leq \prod_{m=0}^{\nu} P_{\boldsymbol{h}} \left\{ |h_m|^2 < \frac{2^R - 1}{\mathsf{SNR}(1 - 2^{R - R_{\max}})} \right\} \doteq \mathsf{SNR}^{-(\nu+1)} . \tag{24}
$$

Therefore, for the spectral efficiencies $R \in (0, R_{\max})$ we have $P_{\mathrm{out}}^{\mathrm{mmse}}(R, \mathsf{SNR}) \overset{\cdot}{\leq} \mathsf{SNR}^{-(\nu+1)}$, which in conjunction with (11) proves that $P_{\mathrm{err}}^{\mathrm{mmse}} \overset{\cdot}{\leq} \mathsf{SNR}^{-(\nu+1)}$, indicating that a diversity order of at least $(\nu + 1)$ is achievable. On the other hand, since the diversity order cannot exceed the number of the CIR taps, the achievable diversity order is exactly $(\nu + 1)$. Also note that the real number $C > 0$ given in

(19) is a constant independent of the CIR realization $\boldsymbol{h}$ and, therefore, C' and, consequently, R_{max} are also independent of the CIR realization. This establishes the proof of Theorem 1 for the range of the spectral efficiencies $R \in (0, R_{\mathrm{max}})$, where R_{max} is fixed and defined in (22).

B. Full Diversity for All Rates

We now extend the results previously found for $R < R_{\mathrm{max}}$ to all spectral efficiencies.

Lemma 2: For asymptotically large values of SNR, $\gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h})$ varies linearly with SNR, i.e.,

$$\lim_{\mathsf{SNR} \to \infty} \frac{\partial \, \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h})}{\partial \, \mathsf{SNR}} = s(\boldsymbol{h}), \quad \text{where} \quad s(\boldsymbol{h}) : \mathbb{R}^{\nu+1} \to \mathbb{R} \, .$$

Proof: See Appendix A. ∎

Lemma 3: For the continuous random variable X, variable $y \in \mathbb{R}$, constants $c_1, c_2 \in \mathbb{R}$ and function $G(X, y)$ continuous in y, we have

$$\lim_{y \to y_0} P_X \Big(c_1 \leq G(X, y) \leq c_2 \Big) = P_X \Big(c_1 \leq \lim_{y \to y_0} G(X, y) \leq c_2 \Big) \, .$$

Proof: Follows from Lebesgue's Dominated Convergence theorem [11] and the same line of argument as in [5, Appendix C] ∎

Now, we show that if for some spectral efficiency $R^{\dagger}$ the achievable diversity order is d, then for all spectral efficiencies *up* to $R^{\dagger} + 1$, the same diversity order is achievable. By induction, we conclude that the diversity order remains unchanged by changing the data spectral efficiency R. If for the spectral efficiency $R^{\dagger}$, the negative of the exponential order of the outage probability is d, i.e.,

$$P_{\boldsymbol{h}} \Big(\log \big[1 + \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h}) \big] < R^{\dagger} \Big) \doteq \mathsf{SNR}^{-d}, \tag{25}$$

then by applying the results of Lemmas 2 and 3 for the target spectral efficiency $R^{\dagger} + 1$ we get

$$P_{\mathrm{out}}^{\mathrm{mmse}}(R, \mathsf{SNR}) = P_{\boldsymbol{h}} \Big(\log \big[1 + \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h}) \big] < R^{\dagger} + 1 \Big) = P_{\boldsymbol{h}} \Big(1 + \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h}) < 2^{R^{\dagger}+1} \Big)$$

$$\doteq P_{\boldsymbol{h}} \Big(\mathsf{SNR} \, s(\boldsymbol{h}) < 2^{R^{\dagger}+1} \Big) = P_{\boldsymbol{h}} \Big(\Big(\frac{\mathsf{SNR}}{2} \Big) s(\boldsymbol{h}) < 2^{R^{\dagger}} \Big) \tag{26}$$

$$\doteq P_{\boldsymbol{h}} \Big(1 + \gamma_{\mathrm{mmse}} \Big(\frac{\mathsf{SNR}}{2}, \boldsymbol{h} \Big) < 2^{R^{\dagger}} \Big) \doteq P_{\boldsymbol{h}} \Big(\log \big[1 + \gamma_{\mathrm{mmse}} \Big(\frac{\mathsf{SNR}}{2}, \boldsymbol{h} \Big) \big] < R^{\dagger} \Big) \tag{27}$$

$$\doteq \Big(\frac{\mathsf{SNR}}{2} \Big)^{-d} \doteq \mathsf{SNR}^{-d} \, . \tag{28}$$

Equations (26) and (27) are derived as the immediate results of Lemmas 2 and 3 that enable interchanging the probability and the limit and also show that $\gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h}) \doteq \mathsf{SNR} \cdot s(\boldsymbol{h})$. Equations (25)-(28) imply that the diversity orders achieved for the spectral efficiencies up to $R^{\dagger}$ and the spectral efficiencies up to $R^{\dagger} + 1$ are the same. As a result, any arbitrary spectral efficiency exceeding R_{max} achieves the same spectral efficiency as the spectral efficiencies $R \in (0, R_{\mathrm{max}})$ and, therefore, for any arbitrary spectral efficiency R, full diversity is achievable via MMSE linear equalization which completes the proof. Figure

1 depicts our simulation results for the pairwise error probabilities for two ISI channels with memory lengths $\nu = 1$ and 2 and MMSE equalization. For each of these channels we consider signal transmission with spectral efficiencies $R = (1, 2, 3, 4)$ bits/sec/Hz. The simulation results confirm that for a channel with two taps the achievable diversity order is two irrespective of the data spectral efficiency. Similarly, it is observed that for a three-tap channel the achievable diversity order is three.

IV. DIVERSITY ORDER OF ZF LINEAR EQUALIZATION

In this section, we show that the diversity order achieved by zero-forcing linear equalization, unlike that achievable with MMSE equalization, is independent of the channel memory length and is always 1.

Lemma 4: For any arbitrary set of normal complex Gaussian random variables $\boldsymbol{\mu} \triangleq (\mu_1, \ldots, \mu_m)$ (possibly correlated) and for any $B \in \mathbb{R}^+$ we have

$$P_{\boldsymbol{\mu}}\left(\sum_{k=1}^{m} \frac{1}{\mathsf{SNR}|\mu_k|^2} > B \right) \dot{\geq} \mathsf{SNR}^{-1} . \tag{29}$$

Proof: Define $W_k \triangleq -\frac{\log |\mu_k|^2}{\log \mathsf{SNR}}$. Since $|\mu_k|^2$ has exponential distribution, it can be shown that for any k the cumulative density function (CDF) at the asymptote of high values of SNR satisfies [12]

$$1 - F_{W_k}(w) \doteq \mathsf{SNR}^{-w} . \tag{30}$$

Thus, by substituting $|\mu_k|^2 \doteq \mathsf{SNR}^{1-W_k}$ based on (30) we find that

$$P_{\boldsymbol{\mu}}\left(\sum_{k=1}^{m} \frac{1}{\mathsf{SNR}|\mu_k|^2} > B \right) \doteq P_{\boldsymbol{W}}\left(\sum_{k=1}^{m} \mathsf{SNR}^{W_k-1} > B \right) \doteq P_{\boldsymbol{W}}(\max_k W_k > 1) \tag{31}$$

$$\geq P_{W_k}(W_k > 1) = 1 - F_{W_k}(1) \doteq \mathsf{SNR}^{-1} . \tag{32}$$

Equation (31) holds as the term $\mathsf{SNR}^{\max W_k - 1}$ is the dominant term in the summation $\sum_{k=1}^{m} \mathsf{SNR}^{W_k-1}$. Also, the transition from (31) to (32) is justified by noting that $\max_k W_k \geq W_k$ and the last step is derived by taking into account (30). ∎

Theorem 3: The diversity order achieved by symbol-by-symbol ZF linear equalization is one, i.e.,

$$P_{\text{err}}^{\text{zf}}(R, \mathsf{SNR}) \doteq \mathsf{SNR}^{-1}$$

Proof: By recalling the decision-point signal-to-noise ratio of ZF equalization given in (5) we have

$$P_{\text{out}}^{\text{zf}}(R, \mathsf{SNR}) = P_{\boldsymbol{h}}\left(\gamma_{\text{zf}}(\mathsf{SNR}, \boldsymbol{h}) < 2^R - 1 \right) = P_{\boldsymbol{h}}\left\{ \left[\frac{1}{2\pi} \int_{\pi}^{\pi} \frac{1}{\mathsf{SNR}|H(e^{-ju})|^2} \, du \right]^{-1} < 2^R - 1 \right\} \tag{33}$$

$$= P_{\boldsymbol{h}}\left\{ \lim_{\Delta \to 0} \left[\sum_{k=0}^{\lfloor 2\pi/\Delta \rfloor} \frac{\Delta}{\mathsf{SNR}|H(e^{-j(-\pi+k\Delta)})|^2} \right]^{-1} < \frac{2^R - 1}{2\pi} \right\} \tag{34}$$

$$= \lim_{\Delta \to 0} P_{\boldsymbol{h}}\left\{ \left[\sum_{k=0}^{\lfloor 2\pi/\Delta \rfloor} \frac{\Delta}{\mathsf{SNR}|H(e^{-j(-\pi+k\Delta)})|^2} \right]^{-1} < \frac{2^R - 1}{2\pi} \right\} \tag{35}$$

$$= \lim_{\Delta \to 0} P_{\boldsymbol{h}} \left\{ \sum_{k=0}^{\lfloor 2\pi/\Delta \rfloor} \frac{\Delta}{\mathsf{SNR}|H(e^{-j(-\pi+k\Delta)})|^2} > \frac{2\pi}{2^R - 1} \right\} \doteq \mathsf{SNR}^{-1} . \tag{36}$$

Equation (34) is derived by using Riemann integration, and (35) holds by using Lemma 3 which allows for interchanging the limit and the probability. Equation (36) holds by applying Lemma 4 on $\mu_k = H(e^{-j(-\pi+k\Delta)})$ which can be readily verified to have Gaussian distribution. Therefore, the achievable diversity order is 1. ∎

Figure 2 illustrates the pairwise error probability of two ISI channels with memory lengths $\nu = 1$ and 2. The simulation results corroborate our analysis showing that the achievable diversity order is one, irrespective of the channel memory length or communication spectral efficiency.

V. CONCLUSION

We showed that infinite-length symbol-by-symbol MMSE linear equalization can fully capture the underlying frequency diversity of the ISI channel. Specifically, the diversity order achieved is equal to that of MLSD and in the high-SNR regime, the performance of MMSE linear equalization and MLSD do not differ in diversity gain and the origin of their performance discrepancy is their ability to control the residual inter-symbol interference. We also show that the diversity order achieved by symbol-by-symbol ZF linear equalizers is always one, regardless of channel memory length.

APPENDIX A
PROOF OF LEMMA 2

We define $g(\boldsymbol{h}, u) \triangleq |H(e^{-ju})|^2$ which has a finite number of zero by following the same line as for $f(\boldsymbol{h}, u)$ in the proof of Lemma 1. By using (8) we get

$$\frac{\partial \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h})}{\partial \mathsf{SNR}} = \frac{\partial}{\partial \mathsf{SNR}} \left(\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\mathsf{SNR}g(\boldsymbol{h}, u) + 1} \, du \right]^{-1} - 1 \right)$$

$$= \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(\boldsymbol{h}, u)}{(\mathsf{SNR}g(\boldsymbol{h}, u) + 1)^2} \, du \right] \cdot \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\mathsf{SNR}g(\boldsymbol{h}, u) + 1} \, du \right]^{-2} \tag{37}$$

$$= \left[\frac{1}{2\pi} \int_{g(\boldsymbol{h}, u) \neq 0} \frac{g(\boldsymbol{h}, u)}{(\mathsf{SNR}g(\boldsymbol{h}, u) + 1)^2} \, du \right] \cdot \left[\frac{1}{2\pi} \int_{g(\boldsymbol{h}, u) \neq 0} \frac{1}{\mathsf{SNR}g(\boldsymbol{h}, u) + 1} \, du \right]^{-2}, \tag{38}$$

where (38) was obtained by removing a finite-number of points from the integral in (37).

Theorem 4: Monotone Convergence [11, Theorem. 4.6]: if a function $F(u, v)$ defined on $U \times [a, b] \to \mathbb{R}$, is positive and monotonically increasing in v, and there exists an integrable function $\hat{F}(u)$, such that $\lim_{v \to \infty} F(u, v) = \hat{F}(u)$, then

$$\lim_{v \to \infty} \int_U F(u, v) \, du = \int_U \lim_{v \to \infty} F(u, v) \, du = \int_U \hat{F}(u) \, du. \tag{39}$$

For further simplifying (38), we define $F_1(u, \mathsf{SNR})$ and $F_2(u, \mathsf{SNR})$ over $\left\{ u \mid u \in [-\pi, \pi], g(\boldsymbol{h}, u) \neq 0 \right\} \times [1, +\infty]$ as

$$
\begin{aligned}
F_1(u, \mathsf{SNR}) &\triangleq \frac{1}{g(\boldsymbol{h}, u)} - \frac{1}{\mathsf{SNR}^2 g(\boldsymbol{h}, u)} + \frac{g(\boldsymbol{h}, u)}{(\mathsf{SNR} g(\boldsymbol{h}, u) + 1)^2}, \\
\text{and} \quad F_2(u, \mathsf{SNR}) &\triangleq \frac{1}{g(\boldsymbol{h}, u)} - \frac{1}{\mathsf{SNR} g(\boldsymbol{h}, u)} + \frac{1}{\mathsf{SNR} g(\boldsymbol{h}, u) + 1}.
\end{aligned}
$$

It can be readily verified that $F_i(u, \mathsf{SNR}) > 0$ and $F_i(u, \mathsf{SNR})$ is increasing in SNR. Moreover, there exist $\hat{F}(u)$ such that

$$
\hat{F}(u) = \lim_{\mathsf{SNR} \to \infty} F_1(u, \mathsf{SNR}) = \lim_{\mathsf{SNR} \to \infty} F_2(u, \mathsf{SNR}) = \frac{1}{g(\boldsymbol{h}, u)}.
$$

Therefore, by exploiting the result of Theorem 4 we find

$$
\lim_{\mathsf{SNR} \to \infty} \int \left[\frac{1}{g(\boldsymbol{h}, u)} - \frac{1}{\mathsf{SNR}^2 g(\boldsymbol{h}, u)} + \frac{g(\boldsymbol{h}, u)}{(\mathsf{SNR} g(\boldsymbol{h}, u) + 1)^2} \right] du = \int \frac{1}{g(\boldsymbol{h}, u)} \, du,
$$

$$
\text{and} \quad \lim_{\mathsf{SNR} \to \infty} \int \left[\frac{1}{g(\boldsymbol{h}, u)} - \frac{1}{\mathsf{SNR} g(\boldsymbol{h}, u)} + \frac{1}{\mathsf{SNR} g(\boldsymbol{h}, u) + 1} \right] du = \int \frac{1}{g(\boldsymbol{h}, u)} \, du,
$$

or equivalently,

$$
\lim_{\mathsf{SNR} \to \infty} \frac{1}{2\pi} \int \frac{g(\boldsymbol{h}, u) \, du}{(\mathsf{SNR} g(\boldsymbol{h}, u) + 1)^2} = \lim_{\mathsf{SNR} \to \infty} \frac{1}{2\pi} \int \frac{du}{\mathsf{SNR}^2 g(\boldsymbol{h}, u)}, \tag{40}
$$

$$
\text{and} \quad \lim_{\mathsf{SNR} \to \infty} \frac{1}{2\pi} \int \frac{du}{\mathsf{SNR} g(\boldsymbol{h}, u) + 1} = \lim_{\mathsf{SNR} \to \infty} \frac{1}{2\pi} \int \frac{du}{\mathsf{SNR} g(\boldsymbol{h}, u)}. \tag{41}
$$

By using the equalities in (40)-(41) and proper replacement in (38) we get

$$
\begin{aligned}
\lim_{\mathsf{SNR} \to \infty} & \frac{\partial \, \gamma_{\mathrm{mmse}}(\mathsf{SNR}, \boldsymbol{h})}{\partial \, \mathsf{SNR}} \\
&= \lim_{\mathsf{SNR} \to \infty} \left[\frac{1}{2\pi} \int_{g(\boldsymbol{h}, u) \neq 0} \frac{g(\boldsymbol{h}, u)}{(\mathsf{SNR} g(\boldsymbol{h}, u) + 1)^2} \, du \right] \cdot \left[\frac{1}{2\pi} \int_{g(\boldsymbol{h}, u) \neq 0} \frac{1}{\mathsf{SNR} g(\boldsymbol{h}, u) + 1} \, du \right]^{-2} \quad (42) \\
&= \lim_{\mathsf{SNR} \to \infty} \left[\frac{1}{2\pi} \int_{g(\boldsymbol{h}, u) \neq 0} \frac{1}{\mathsf{SNR}^2 g(\boldsymbol{h}, u)} \, du \right] \cdot \left[\frac{1}{2\pi} \int_{g(\boldsymbol{h}, u) \neq 0} \frac{1}{\mathsf{SNR} g(\boldsymbol{h}, u)} \, du \right]^{-2} \\
&= \left[\frac{1}{2\pi} \int_{g(\boldsymbol{h}, u) \neq 0} \frac{1}{g(\boldsymbol{h}, u)} \, du \right]^{-1} = s(\boldsymbol{h}),
\end{aligned}
$$

where $s(\boldsymbol{h})$ is independent of SNR and thus the proof is completed.

REFERENCES

[1] J. Proakis, *Digital Communications*, 3rd ed.   McGraw Hill, 1995.

[2] J. G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.

[3] S. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, vol. 9, no. 73, pp. 1349– 1387, September 1985.

[4] D. P. Taylor, G. M. Vitetta, B. D. Hart, and A. Mmmel, "Wireless channel equalization," *Eur. Trans. Telecommun*, vol. 9, no. 2, pp. 117–143, 1998.
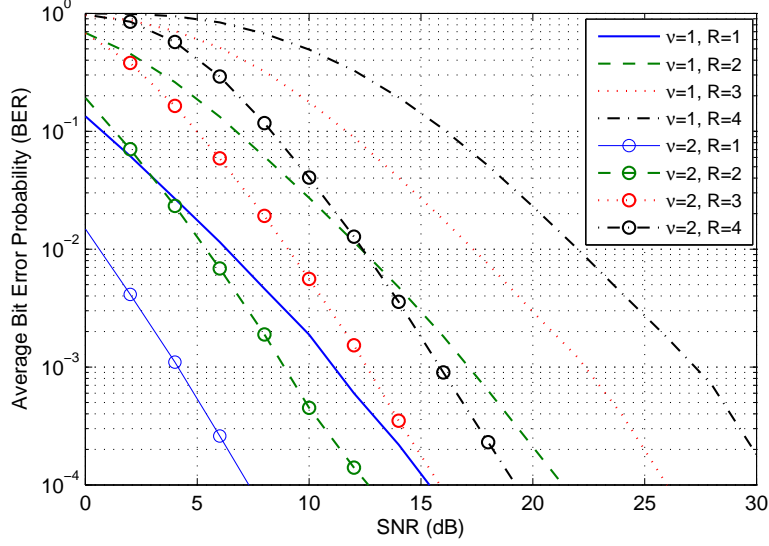
Fig. 1. Average detection error probability in two-tap and three-tap ISI channels with MMSE linear equalization.

[5] A. Tajer, A. Nosratinia, and N. Al-Dhahir, "MMSE infinite length symbol-by-symbol linear equalization achieves full diversity," in *Proc. IEEE International Symposium on Information Theory (ISIT'07)*, Nice, France, June 2007, pp. 1706–1710.

[6] J. M. Cioffi, *Digital Communication: Signal Processing*, http://www.stanford.edu/class/ee379a/course_reader/chap3.pdf.

[7] J. M. Cioffi, G. Dudevoir, M. Eyuboglu, and J. G.D. Forney, "Minimum mean-square-error decision feedback equalization and coding - parts I and II," *IEEE Trans. Commun.*, vol. 43, no. 10, pp. 2582–2604, October 1998.

[8] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels,," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[9] M. J. D. Powell, *Approximation Theory and Methods*. Cambridge University Press, 1981.

[10] O. C. McGehee, L. Pigno, and B. Smith, "Hardy's inequality and the ℓ_1 norm of exponential sums," *Anals of Matehmatics*, vol. 113, pp. 613–618, 1981.

[11] R. G. Bartle, *The Elements of Integrations*. John Wiley & Sons Inc, 1966.

[12] K. Azarian, H. El-Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.
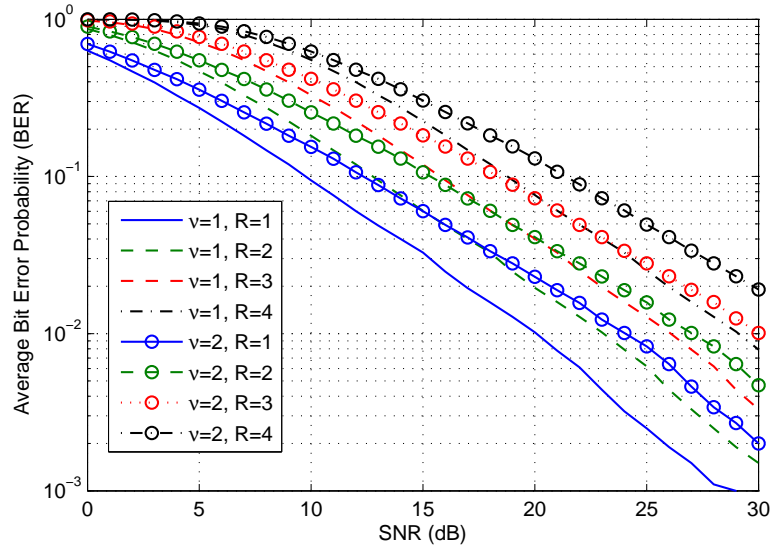
Fig. 2. Average detection error probability in two-tap and three-tap ISI channels with ZF linear equalization